An Overview of Wavelet Analysis and Time-Frequency Analysis (a minicourse)

Bruno Torrésani

An Overview of Wavelet Analysis and Time-Frequency Analysis (a minicourse)

Bruno Torrésani[1]

LATP, CMI, Université de Provence,
39 Avenue F. Joliot-Curie, 13432 Marseille
FRANCE

Abstract

We describe several aspects of wavelet analysis and more general methods of time-frequency analysis, emphasizing applications to signal analysis and processing problems.

1 Introduction

Time-Frequency analysis has been introduced from the need of developing a mathematical version of musical notation. Classical analysis is generally based upon a representation of functions, either the usual one (hereafter termed the "time representation") or the Fourier representation ("frequency representation"): formally

$$\hat{f}(\omega) = \int f(t) e^{-i\omega t} dt \ .$$

It is known that the frequency representation yields a lot of difficult mathematical problems (see e.g. [57]). However, the frequency representation has also been criticized for its lack of physical significance. From a physical point of view, it does not make sense to think of a function or a signal as being a superposition of functions which do not possess any time localization properties. Let us take the example of a musical signal. Our ear interprets it as a series of notes, or "atoms of sound" appearing at given times, with a finite duration and a given height (the fundamental frequency). If the signal contains a given note once, say A, a Fourier representation of the signal will exhibit a peak at the corresponding frequency, without any indication of location or duration. We now quote J. Ville:

> *"...the representation is mathematically correct because the phases of the tones close to A have managed to suppress it by interference phenomena before it is heard, and to enforce it, again by interferences, when it is heard... However this is a deformation of reality: when the A is not heard, it is simply because it has not been played yet..."*

This motivates the search for a mathematical analogue of the musical notation: a way of representing functions in terms of joint time-frequency variables.

[1]email address: `Bruno.Torresani@sophia.inria.fr`

This programme of *time-frequency representations* has been achieved in various ways and various contexts. The first contributions seem to be due to Ville on one hand, who proposed to use the Wigner distribution as a "time-frequency density", and Gabor on the other hand, whose approach was based upon decompositions into "time-frequency atoms", generated as time and frequency translates of a Gaussian function. Since then, Ville's approach has led to the theory of quadratic time-frequency representations, very popular in the signal analysis community. Gabor-type expansions are also very popular, and have more recently enjoyed a renewed interest, since the development of wavelet theory in the early eighties.

A. Grossmann and J. Morlet realized in 1983 that any square-integrable function may be expanded into wavelets, namely translates and dilates of a single function (the mother wavelet), provided that the latter possess some mild oscillation properties. This simple result, and the remark that the corresponding expansions are in fact very close to Gabor's expansions, showed that wavelet expansions had an important potential as a tool for signal analysis. Grossmann and his collaborators also provided a beautiful interpretation of wavelet and Gabor expansions in terms of square-integrable group representations, emphasizing the importance of symmetry groups in the construction.

At about the same time, motivated by image processing problems, and the construction by Y. Meyer of an orthonormal basis of $L^2(\mathbb{R})$ consisting in a discrete system of wavelets (see [46] for example), S. Mallat [42] developed the theory of *Multiresolution analysis*, providing a sort of algorithmic framework for generating wavelet bases. This discovery had in fact a tremendous impact in various fields, as it realized a synthesis between several apparently disconnected fields: image analysis and computer vision, mathematical analysis, signal coding,... A major achievement was the construction by I. Daubechies [16] of orthonormal bases of compactly supported wavelets. The construction was generalized in many respects since then: construction of new compactly supported wavelet orthonormal bases (see [15] for a general account), biorthogonal bases [13], wavelet bases on intervals and open sets, generalized subband decompositions [14, 55] (wavelet packets, local trigonometric bases,...), and many other developments.

The wavelet theory has also evolved very much along lines motivated by applications. Besides signal and image processing applications, which were in some sense the first motivations for wavelet-based methods, new applications have gradually emerged. Without being exhaustive again, let us quote applications to numerical analysis [7], dynamical systems [4] and turbulence, quantum physics, and more recently statistics. All these subjects brought additional contributions to the general theory of wavelets.

The present paper is a written version of lectures given by the author at the Dubna workshop on "self-similar systems", in July 1998. The goal of the lectures was to present a general introduction to the wavelet theory and some more recent developments, with a necessarily "coarse sampling" of the subject. Obviously, the theory is by now too large to fit into a single survey paper. Therefore, we have chosen to limit our presentation to specific aspects, mostly related to signal processing problems. We start our presentation with the introduction of wavelet transforms in the context of time-frequency analysis, in parallel with the so-called "quadratic time-frequency transforms", and the continuous

Gabor transform. After describing some classical time-frequency problems, we illustrate the

2 Time-frequency analysis; continuous transforms

The whole story of wavelet analysis started with the introduction of the continuous wavelet transform in a seminal article by A. Grossmann and J. Morlet [29]. Many earlier contributions should in fact be quoted, both in the signal processing and mathematical literature, but it was only in that paper that wavelet transform was considered as such.

Like the continuous Gabor transform (also called short-time Fourier transform), the continuous wavelet transform is a prototype of the so-called *linear time frequency representations*, which provide decompositions of functions as superpositions of elementary waveforms. Those linear time-frequency representations are often compared with the *quadratic time-frequency representations*, a family of tools which are considered more powerful in some respects, but generally lack of robustness in complex situations. We start our analysis by a short account of the quadratic time-frequency representations.

2.1 Quadratic time-frequency transforms

Definitions and general properties The first instance of time-frequency transform is the *ambiguity function*, familiar to radar specialists (see e.g. [8]). The (cross) ambiguity function $\mathcal{A}_{f,g}(b,\omega)$ of a pair of functions $f(t)$ and $g(t)$ may be introduced as the solution of the following problem: let $g(t)$ be a fixed reference signal, and let $f(t)$ be an observation, assumed to be a noisy version of a time and frequency shifted copy of $g(t)$:

$$f(t) = Ag(t-\tau)e^{i\xi t} + n(t) \ ,$$

where A is a constant and $n(t)$ is a random perturbation, supposed to be a white noise for the sake of simplicity. Let us assume that $g \in L^2(\mathbb{R})$. If the random component is absent, the problem may be formulated as follows: find a family of continuous linear forms $L_{b,\omega}$ such that for all $\tau, \xi)$, $L_{b,\omega}(g_{\tau,\xi})$ attains its maximum for $(b,\omega) = (\tau,\xi)$. If $n(t)$ is present, $L_{b,\omega}(f) = AL_{b,\omega}(g_{\tau,\xi}) + L_{b,\omega}(n)$. According to Riesz's representation theorem, there exists a function $\varphi_{b,\omega}(t) \in L^2(\mathbb{R})$ such that for all $f \in L^2(\mathbb{R})$, $L_{b,\omega}(f) = \langle f, \varphi_{b,\omega}\rangle$. Then, denoting by $\mathbb{E}\{X\}$ the expectation of a random variable X, $\mathbb{E}\left\{|L_{b,\omega}(n)|^2\right\} = ||\varphi_{b,\omega}||^2$, and the optimal $\varphi_{b,\omega}$ is defined as the one which maximizes the *Signal to Noise Ratio* (SNR) $\rho^2 = |\langle g_{\tau,\xi}, \varphi_{\tau,\xi}\rangle|^2/\mathbb{E}\left\{|L_{b,\omega}(n)|^2\right\}$. The solution $\varphi_{b,\omega} = Kg_{b,\omega}$ (where $K \neq 0$ is a constant) is an immediate consequence of the Cauchy-Schwarz inequality. Therefore, the parameters τ and ξ may be obtained by maximizing with respect to b,ω the square-modulus of

$$\int f(t)\overline{g}(t-b)e^{-i\omega t}dt \ .$$

The latter is essentially a time-frequency cross-correlation of $f(t)$ and $g(t)$, obtained by considering scalar products of $f(t)$ with time and frequency shifted copies of $g(t)$ of the form $e^{i\omega t}g(t-b)$ (with $b,\omega \in \mathbb{R}$). Therefore, it measures how close $f(t)$ is to time-frequency shifted copies of $g(t)$.

The definition (including the case of random time series) is given below.

Definition 2.1 *1. Let $f \in L^2(\mathbb{R})$. Its ambiguity function is defined by*

$$\mathcal{A}_f(\tau, \xi) = \int f(t + \tau/2)\overline{f}(t - \tau/2)e^{-i\xi t}dt \ . \tag{1}$$

2. Let $\{X_t, t \in \mathbb{R}\}$ a second order random time series. Then its ambiguity function is defined by

$$\mathcal{A}_X(\tau, \xi) = \mathbb{E}\left\{\int X_{t+\tau/2}\overline{X}_{t-\tau/2}e^{i\xi t}dt\right\} \ . \tag{2}$$

It is easily seen that if $f \in L^2(\mathbb{R})$, then $\mathcal{A}_f$ is a bounded function (with $||\mathcal{A}_f||_\infty \le ||f||^2$). In addition, a direct calculation shows that if $f \in L^2(\mathbb{R})$, then $\mathcal{A}_f \in L^2(\mathbb{R}^2)$, and that $||\mathcal{A}_f||_2^2 = 2\pi||f||^4$. More generally, it may be shown that $\mathcal{A} \in L^p(\mathbb{R}^2)$ for all $p \in [1, \infty]$ (bounds for the corresponding $L^p(\mathbb{R}^2)$ norms have been derived by E. Lieb [39]). More generally, the definition of ambiguity function may be extended to distributions, and it may be shown that the "ambiguity distribution" of a distribution $\phi \in \mathcal{S}'(\mathbb{R})$ is a distribution $\mathcal{A}_\phi \in \mathcal{S}'(\mathbb{R}^2)$.

The ambiguity and cross-ambiguity functions possess several interesting properties. Among them, the so-called Moyal's formula is extremely important:

Proposition 2.1 *Let $f, f', g, g' \in L^2(\mathbb{R})$. Then $\mathcal{A}_{f,g}, \mathcal{A}_{f',g'} \in L^2(\mathbb{R}^2)$, and*

$$\langle \mathcal{A}_{f,g}, \mathcal{A}_{f',g'}\rangle = 2\pi\langle f, f'\rangle\,\overline{\langle g, g'\rangle} \ . \tag{3}$$

The proof of the proposition follows from a simple calculation.

The non-deterministic version may be given a similar interpretation. Its properties depend on the properties of the covariance operator $\mathcal{C}$ of the process, defined by its matrix elements: for all $f, g \in \mathcal{D}(\mathbb{R})$,

$$\langle \mathcal{C}f, g\rangle = \mathbb{E}\left\{\langle X, g\rangle\overline{\langle X, f\rangle}\right\} \ . \tag{4}$$

For example, if $\mathcal{C}$ extends to a Hilbert-Schmidt operator, which we denote by $\mathcal{C} \in \mathcal{L}^2$, then $\mathcal{A} \in L^2(\mathbb{R}^2)$.

A convenient way of expressing the ambiguity function makes use of the translation operator T_b, defined by $T_b f(t) = f(t - b)$ and the modulation operator E_ω defined by $E_\omega f(t) = e^{i\omega t}f(t)$: $\mathcal{A}_{f,g}(\tau, \xi) = \langle T_{-\frac{\tau}{2}}E_{-\frac{\xi}{2}}f, T_{\frac{\tau}{2}}E_{\frac{\xi}{2}}g\rangle$. This shows in particular that the ambiguity function $\mathcal{A}_f(\tau, \xi)$ only provide estimates for the *spreading* of the analyzed object in the joint time-frequency plane, but not on its localization in that space. Such an analysis is done by the *Wigner-Ville* (WV) distribution, defined by

Definition 2.2 *1. Let $f \in L^2(\mathbb{R})$. Its Wigner-Ville distribution is the function of two real variables $\mathcal{W}_f$ defined by*

$$\mathcal{W}_f(b, \omega) = \int f(b + \tau/2)\overline{f}(b - \tau/2)e^{-i\omega\tau}d\tau \tag{5}$$

2. Let $\{X_t, t \in \mathbb{R}\}$ a second order time series. Then its Wigner-Ville distribution (or Wigner spectrum) is defined by

$$\mathcal{E}_X(b, \omega) = \mathbb{E}\left\{\int X_{b+\tau/2}\overline{X}_{b-\tau/2}e^{-i\omega\tau}d\tau\right\} \tag{6}$$

Remark 2.1 It is readily seen that the Wigner-Ville function and the ambiguity function are related via a symplectic Fourier transform

$$
\begin{aligned}
\mathcal{A}(\tau,\xi) &= \frac{1}{2\pi}\int\int \mathcal{W}(b,\omega)e^{-i(\xi b-\omega\tau)}db\,d\omega\ , &(7)\\
\mathcal{W}(b,\omega) &= \frac{1}{2\pi}\int\int \mathcal{A}(\tau,\xi)e^{i(\xi b-\omega\tau)}d\tau\,d\xi\ . &(8)
\end{aligned}
$$

The same holds true in the non-deterministic context. Therefore, the Wigner function is square-integrable as soon as the ambiguity function is so. Therefore, $\mathcal{W}_f \in L^2(\mathbb{R}^2)$ as soon as $f \in L^2(\mathbb{R})$. In fact, Lieb's estimates show that when $f \in L^2(\mathbb{R})$, $\mathcal{W}_f \in L^p(\mathbb{R}^2)$ for all $p \in [1,\infty]$.

Remark 2.2 The Wigner-Ville distribution may also be defined when f is a distribution. Indeed, if $f \in \mathcal{S}'(\mathbb{R})$, it may be shown (see [25] for example) that $\mathcal{W}_f \in \mathcal{S}'(\mathbb{R}^2)$. In fact, the Wigner-Ville distribution is nothing but the Weyl symbol of the operator of orthogonal projection onto f, which may be written:

$$
P_f g(t) = \frac{1}{2\pi||f||^2}\int \mathcal{W}_f\left(\frac{t+b}{2},\omega\right)e^{i\omega(t-b)}g(b)dbd\omega\ .
$$

Properties of the Wigner-Ville distribution The WV distributions enjoy remarkable properties. The first one reflects the orthogonality relations of the Ambiguity functions, and go under the name of *Moyal's formulae*: let $f, f', g, g' \in L^2(\mathbb{R})$. Then $\mathcal{W}_{f,g}, \mathcal{W}_{f',g'} \in L^2(\mathbb{R}^2)$ and

$$
\langle \mathcal{W}_{f,g}, \mathcal{W}_{f',g'}\rangle = 2\pi\langle f,f'\rangle\,\langle g',g\rangle\ . \tag{9}
$$

The orthogonality relations of the Wigner-Ville coefficients are a direct consequence of the corresponding relations for ambiguity functions, and the symplectic Fourier transform formulas (7) and (8).

A main property of the WV distribution is its covariance with respect to a certain number of simple transformations. Namely:

1. *Translations:* if $g(t) = f(t-\tau)$, $\mathcal{W}_g(b,\omega) = \mathcal{W}_f(b-\tau,\omega)$.
2. *Modulations:* if $g(t) = e^{i\lambda t}f(t)$, then $\mathcal{W}_g(b,\omega) = \mathcal{W}_f(b,\omega-\lambda)$.
3. *Rescalings:* if $g(t) = \frac{1}{\sqrt{a}}f\left(\frac{t}{a}\right)$, then $\mathcal{W}_g(b,\omega) = \mathcal{W}_f\left(\frac{b}{a},a\omega\right)$.
4. *Time-frequency rotations:* Let θ be such that $\cos\theta \neq 0$, and set

$$
g_\theta(t) = \frac{1}{\sqrt{\cos\theta}}\frac{1}{2\pi}\int e^{i\tan\theta(t^2+\xi^2)/2}e^{it\xi/\cos\theta}\hat{f}(\xi)d\xi\ .
$$

Then

$$
\mathcal{W}_{g_\theta}(b,\omega) = \mathcal{W}_f(b\cos\theta+\omega\sin\theta, -b\sin\theta+\omega\cos\theta)\ .
$$

More generally, the Wigner transform is covariant under a general group of transformations, called the *metaplectic group*. We refer to [25] for more details.

2.2 Examples

Time-frequency localization The examples that best illustrate the optimal localization properties of the WV distribution are the *pure oscillations*, i.e. the distributionof the form $f(t) = e^{i\lambda t}$. The WV transform of such an $f(t)$ has to be defined as a two-dimensional distribution, and one easily shows that $\mathcal{W}_f(b, \omega) = 2\pi\delta(\omega - \lambda)$.

More generally, such optimal localization properties are preserved by the simple transformations alluded to in the previous section. While the effect of translations, modulations and rescalings are easy to understand, let us stress the role played by time-frequency rotations. It is easily shown that the so-called *linear chirps*, i.e. the distributions with a linearly time-varying frequency $e^{i\lambda t + \alpha t^2/2}$ may be obtained by appropriate translation, modulation, rescaling and time-frequency rotation of a pure oscillation. Therefore, its WV distribution inherits the perfect localization properties from those of the pure oscillations, and one obtains $\mathcal{W}_f(b, \omega) = 2\pi\delta(\omega - (\lambda + \alpha b))$. An example of such a behavior is provided in Fig. 1, which represents the WV distribution of a linear chirp.

Interferences As a quadratic functional of the function $f(t)$, the WV distribution yields interference terms. Namely, let $f_1, f_2 \in L^2(\mathbb{R})$, and let $f = f_1 + f_2$. Then one immediately sees that

$$\mathcal{W}_f(b, \omega) = \mathcal{W}_{f_1}(b, \omega) + \mathcal{W}_{f_2}(b, \omega) + 2\Re\mathcal{W}_{f_1, f_2}(b, \omega)$$

Even in the case where both $\mathcal{W}_{f_1}(b, \omega)$ and $\mathcal{W}_{f_2}(b, \omega)$ are sharply localized in the (b, ω) plane, the "cross term" $2\Re\mathcal{W}_{f_1, f_2}(b, \omega)$ introduces an extra component in the WV transform of f, which degrades the resolution. An example is given in Fig. 2, where we display the WV distribution of the sum of three pulses.

The treatment of such interference terms has received a considerable attention in the signal processing literature during the past 10 years. Interferences terms are generally attenuated by appropriate smoothings of the WV distribution. Examples are provided by the following classes of time-frequency distributions:

1. *The Cohen's class:* let $\Pi \in L^1(\mathbb{R}^2)$ be a fixed function. Associate with it the following time-frequency distribution

$$\rho_f(b, \omega) = \int \Pi(\tau, \xi)\mathcal{W}(b - \tau, \omega - \xi) d\tau d\xi \ . \tag{10}$$

   $\rho_f(b, \omega)$ is expected to be smoother than the Wigner-Ville distribution, and in particular to contain much less important interference terms. Notice that since it is obtained from a WV distribution via a two-dimensional convolution (a translation invariant operation), $\rho_f(b, \omega)$ inherits from $\mathcal{W}(b, \omega)$ the time-frequency translations covariance.

2. *The affine class:* Given a function of two variables $\Pi(b, \omega)$, introduce

$$\rho_f(b, a) = \int \Pi(\tau, \xi)\mathcal{W}\left(\frac{b - \tau}{a}, a\xi\right) d\tau d\xi \ . \tag{11}$$

   $\rho_f(b, a)$ is a *time-scale representation* of f. With the same arguments as before, it may be shown that it inherits from the WV distributions its properties of covariance with respect to translations and rescalings.

The need for linear time-frequency transform The quadratic time-frequency distributions we just described have many interesting properties. We refer to [23] for a detailed analysis of such properties in a signal processing perspective. However, they are often difficult to interpret in practical situations. For that reason, simpler alternatives such as *linear time-frequency transforms* are often prefered. We describe below the simplest two examples of such transforms. We first focus on the case of continuous transforms, and postpone the description of the discretization problem to a subsequent section.

2.3 Continuous Gabor transform (CGT)

Definitions The simplest localized version of Fourier analysis is provided by windowed Fourier transform, whose main idea is to localize the signal first by multiplying it by a smooth and localized window, and then perform a Fourier transform. More precisely, the construction goes as follows. Start from a function $g \in L^2(\mathbb{R})$, such that $||g|| \neq 0$, and associate with it the following family of *Gaborlets*

$$g_{(b,\omega)}(t) = e^{i\omega(t-b)}g(t-b) \ . \tag{12}$$

Definition 2.3 *Let $g \in L^2(\mathbb{R})$ be a window. The continuous Gabor transform of a finite-energy signal $f \in L^2(\mathbb{R})$ is defined by the integral transform*

$$G_f(b,\omega) = \langle f, g_{(b,\omega)}\rangle = \int f(t)\,\overline{g(t-b)}e^{-i\omega(t-b)}\,dt \ . \tag{13}$$

Gaborlets yield decomposition formulas for functions in $L^2(\mathbb{R})$, as follows.

Theorem 2.1 *Let $g \in L^2(\mathbb{R})$ be a non trivial window (i.e. $||g|| \neq 0$.) Then every $f \in L^2(\mathbb{R})$ admits the decomposition*

$$f(t) = \frac{1}{2\pi||g||^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_f(b,\omega)g_{(b,\omega)}(t)dbd\omega \ , \tag{14}$$

where equality holds in the weak $L^2(\mathbb{R})$ sense.

In other words, the mapping $L^2(\mathbb{R}) \ni f \hookrightarrow \frac{1}{||g||\sqrt{2\pi}} G_f \in L^2(\mathbb{R}^2)$ is an isometry between $L^2(\mathbb{R})$ and $L^2(\mathbb{R}^2)$, and the inverse mapping is provided by the adjoint mapping.

The case of random time series

Definition 2.4 *Let $\{X_t, t \in \mathbb{R}\}$ be a mean zero second order random time series, and let $g \in L^2(\mathbb{R})$ be a window. The CGT of X_t is the random time-frequency series defined by*

$$G_X(b,\omega) = \langle X, g_{(b,\omega)}\rangle \tag{15}$$

By definition, the covariance operator $\mathcal{C}$ of the time series is defined by its matrix elements $\langle \mathcal{C}f, g\rangle = \mathbb{E}\left\{\langle X,g\rangle\langle f,X\rangle\right\}$. Obviously,

$$\mathbb{E}\left\{G_f(b,\omega)\overline{G_f}(b',\omega')\right\} = \langle \mathcal{C}g_{(b',\omega')}, g_{(b,\omega)}\rangle \tag{16}$$

The case where the time series $\{X_t, t \in \mathbb{R}\}$ under consideration is (second order) stationary is particularly interesting. By definition, the covariance operator is in such a case a convolution operator, and one immediately sees that

$$\mathbb{E}\left\{G_f(b,\omega)\overline{G_f}(b',\omega)\right\} = \frac{1}{2\pi}\int e^{i\eta(b-b')}\mathcal{E}(\eta)\left|\hat{g}(\eta-\omega)\right|^2 d\eta\ ,$$

where $\mathcal{E}(\eta)$ stands for the spectral density of the time series.

However, the main interest of the CGT lies in its potential for handling non stationary situations (in the stationary case, the covariance operator is a convolution operator, which is perfectly handled by Fourier methods). Particularly interesting is the case of the so-called *locally stationary time series*, which are basically random time series whose covariance operator is "almost diagonal" in an appropriate Gabor representation. Such a situation has been discussed by various authors in various contexts (see e.g. [38, 44, 11, 33]).

examples The CGT and similar tools have been quite popular in the speech processing literature, because of its capability of handling the so-called *locally harmonic signals*, namely signals which may be modeled in the form

$$f(t) = \sum_{k=1}^{K} A_k(t)e^{i\phi_h(t)}\ , \tag{17}$$

where the local amplitudes $a_k(t)$ and the local frequencies $\omega_k(t) = \phi_k'(t)$ are assumed to be slowly varying. Such signals are called locally harmonic when, in addition, the local frequencies are close to be integer multiples of a fundamental frequency (the *pitch frequency*: $\omega_k(t) \approx k\omega_1(t)$. The CGT of such signals have been studied by several authors. For example we have the following estimates

Lemma 2.1 *Let $g \in L^2(\mathbb{R})$ be a window such that $\int |tg(t)|dt < \infty$ and $\int t^2|g(t)|dt < \infty$. Let*

$$f(t) = A(t)e^{i\phi(t)}$$

be such that $A, \phi \in C^2(\mathbb{R})$, and let $G_f(b,\omega)$ denote the continuous Gabor coefficients of f with respect to the window g. Then

$$G_f(b,\omega) = A(b)e^{i\phi(b)}\overline{\hat{g}}(\phi'(b)-\omega)\left[1 + R(b,\omega)\right]\ , \tag{18}$$

and the remainder $R(b,\omega)$ is bounded as

$$|R(b,\omega)| \le K_1\frac{|A'(b)|}{|A(b)|} + K_2\frac{\sup_u|A''(u)|}{|A(b)|} + K_3\frac{\sup_u|\phi''(u)|}{|A(b)|}\ .$$

A proof of a statement close to that one may be found in [41]. See also [19, 51, 11] for similar approaches.

The meaning of such statements is the following: as soon as the amplitudes $A(t)$ and frequencies $\phi'(t)$ are slowly varying enough, one may write

$$G_f(b,\omega) \approx A(b)e^{i\phi(b)}\overline{\hat{g}}(\phi'(b)-\omega)\ .$$

Assume now that $g(t)$ is a smooth function, located near the origin $t = 0$ (typically a Gaussian function.) The latter expression shows that for each value of the time variable b, the modulus of the CGT attains its maximum on a curve (the so-called *ridge*) of equation $\omega = \phi'(b)$, i.e. describing the instantaneous frequency of the function $f(t)$.

We refer to [11] for a large number of numerical illustrations of such a situation. We limit our illustrations to an example of speech signal, in which several such components are present. Thanks to the linearity of the CGT, the CGT of a signal of the type (17) is of the form

$$G_f(b,\omega) \approx \sum_{k=1}^{K} A_k(b) e^{i\phi_k(b)} \overline{\hat{g}}(\phi_k'(b) - \omega) \ ,$$

and as soon as g is such that the various $\phi_k'(b) - \phi_\ell'(b)$ for $\ell \neq k$ are small enough, the CGT of such a signal localizes itself near K different ridges. An example of such a behavior is given in Fig. 3, where we display 625 milliseconds of speech signal: /How are you ?/ (top) and the modulus of the CGT. Observe the localisation near the ridges.

2.4 Continuous Wavelet Transform (CWT)

An alternative to Gabor transform was proposed more recently by Grossmann and Morlet [29]. The main idea was to improve the time resolution of Gabor transform, by changing the rule for generating the "basis functions". This may be done by replacing the modulation operation used to generate Gaborlets by a scaling operation.

Wavelet transform Let $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a fixed function (in fact, it is sufficient to assume $\psi \in L^1(\mathbb{R})$, but for convenience also assume that $\psi \in L^2(\mathbb{R})$. This extra assumption ensures the boundedness of the wavelet transform.) From now on it will be called the *analyzing wavelet*. It is also sometimes called the *mother wavelet* of the analysis. The corresponding family of wavelets is the family $\{\psi_{(b,a)}; b \in \mathbb{R}, a \in \mathbb{R}_+^*\}$ of shifted and scaled copies of ψ defined as follows. If $b \in \mathbb{R}$ and $a \in \mathbb{R}_+^*$ we set:

$$\psi_{(b,a)}(t) = \frac{1}{a}\psi\left(\frac{t-b}{a}\right), \quad t \in \mathbb{R} \ . \tag{19}$$

The wavelet $\psi_{(b,a)}$ can be viewed as a copy of the original wavelet ψ rescaled by a and centered around the "time" b. Given an analyzing wavelet ψ, the associated continuous wavelet transform is defined as follows

Definition 2.5 *Let $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be an analyzing wavelet. The continuous wavelet transform (CWT for short) of a finite-energy signal $f(t)$ is defined by the integral:*

$$T_f(b,a) = \langle f, \psi_{(b,a)} \rangle = \frac{1}{a}\int f(t)\overline{\psi}\left(\frac{t-b}{a}\right) dt \ . \tag{20}$$

Like Gaborlets, wavelets may form complete sets of functions in $L^2(\mathbb{R})$, and we have in particular

Theorem 2.2 *Let $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, be such that the number c_ψ defined by:*

$$c_\psi = \int_0^\infty |\hat{\psi}(a\xi)|^2 \frac{da}{a} \tag{21}$$

is finite, nonzero and independent of $\xi \in \mathbb{R}$. Then every $f \in L^2(\mathbb{R})$ admits the decomposition

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^\infty \int_0^\infty T_f(b,a)\psi_{(b,a)}(t) \frac{da}{a} db \ , \tag{22}$$

where the convergence holds in the strong $L^2(\mathbb{R})$ sense.

In particular, we also have "energy conservation": if $f \in L^2(\mathbb{R})$, then $T_f \in L^2(\mathbb{R} \times \mathbb{R}_+^*, db\frac{da}{a})$, and $||T_f||^2 = c_\psi ||f||^2$. Notice that in (21), the constant c_ψ can only depend on the sign of $\xi \in \mathbb{R}$, therefore assuming independence wrt ξ is a simple symmetry assumption. The fact that $0 < c_\psi < \infty$ implies that $\hat{\psi}(0) = 0$, so that the wavelet $\psi(t)$ has to oscillate enough to be of zero mean.

The CWT of random time series may be defined in a way similar to the CGT.

Examples The CWT has a behavior similar to that of the CGT in many respects. The main difference lies in the fact that wavelets are extremely precise at small scales (where they lose frequency resolution), and more frequency localized at large scales (where time resolution is lost). A main application of this fact is the analysis of regularity (see below). A visualization of this effect may be found in Fig. 3 (bottom), where we display the modulus of the CWT of the speech signal shown at the top of the figure. As may be seen, at large scales, the wavelets have a sufficient frequency resolution to analyze carefully the first harmonics (namely, the pitch frequency and the first harmonic). For smaller scales, frequency resolution is lost, and the same wavelet is unable to "separate" several harmonic components. This results in interferences between the harmonic components, which yield the oscillations of the modulus in the b direction that appear on the image. We refer to [11] for a more detailed analysis of such applications.

2.5 Linear transforms as approximations

Interestingly enough, there is a strong connection between the Wigner-Ville distributions and the linear decompositions we just reviewed (or more precisely their squared modulus). We express this in the following

Proposition 2.2 *Let $f \in L^2(\mathbb{R})$, and let T_f, G_f and $\mathcal{E}_f$ denote respectively its CWT, CGT and WV transforms. Then the following two properties are true:*

1. *The squared modulus $|G_f(b,\omega)|^2$ of the CGT is a smoothed version of the WV distribution $\mathcal{W}_f$ of f:*

$$|G_f(b,\omega)|^2 = \int \mathcal{W}_f(\tau,\xi)\mathcal{W}_g(\tau - b, \xi - \omega) d\tau d\xi \tag{23}$$

2. *The squared modulus* $|T_f(b,a)|^2$ *of the CWT is an* affine smoothing *of the WV distribution* $\mathcal{W}_f$ *of* f*:*

$$|T_f(b,a)|^2 = \int \mathcal{W}_f(\tau,\xi)\mathcal{W}_\psi\left(\frac{\tau-b}{a}, a\xi\right) d\tau d\xi \qquad (24)$$

These results follow from an easy calculation. See e.g. [23] for more details.

2.6 Littlewood-Paley transform

We close this section with a short account of a simplified version of wavelets decomposition, known for a long time under the name of *Littlewood-Paley decomposition*. Littlewood-Paley decomposition also provide a first step towards the discretization of the continuous wavelet transforms, by a discretization of the scale variable only.

Let $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be such that $0 < c_\psi < \infty$, and let us assume that ϕ is such that

$$\left|\hat\phi(a\xi)\right|^2 = \int_a^\infty \left|\hat\psi(u\xi)\right|^2 \frac{du}{u} \ . \qquad (25)$$

Assuming an obvious definition for $\phi_{(b,a)}(t)$, we set

$$S_f(b,a) = \langle f, \phi_{(b,a)}\rangle \qquad (26)$$

for all $f \in L^2(\mathbb{R})$. Then equation (22) may be replaced by

$$\begin{aligned} f(t) &= \frac{1}{c_\psi} \lim_{a\to 0} \int S_f(b,a)\phi_{(b,a)}(t)db \\ &= \frac{1}{c_\psi}\left(\int S_f(b,a)\phi_{(b,a)}(t)db + \int_{-\infty}^{\infty}\int_0^a T_f(b,u)\psi_{(b,u)}(t)\frac{du}{u}db\right) \ . \end{aligned} \qquad (27)$$

The function ϕ is called a (bilinear) *scaling function* associated with the wavelet ψ. We refer to [22] for more details.

It is important to emphasize the difference between wavelet coefficients T_f and scaling function coefficients S_f. While the coefficients $S_f(b,a)$ provide an approximation of the function f at scale a, $T_f(b,a)$ give details at scale a, i.e., essentially differences between two consecutive approximations.

In fact, the integral with respect to the scale variable in the continuous wavelet decompositions may easily be replaced with a discrete sum, by the following trick. Let $\Psi(t)$ be such that

$$\left|\hat\Psi(\omega)\right|^2 = \int_{1/2}^1 \left|\hat\psi(a\omega)\right|^2 \frac{da}{a} \qquad (28)$$

The, setting for all $f \in L^2(\mathbb{R})$

$$\Sigma_j(f)(t) = \frac{1}{c_\psi}\int S_f(b,2^{-j})\phi_{(b,2^{-j})}(t)db \qquad (29)$$

$$D_f(b,a) = \langle f, \Psi_{(b,a)}\rangle \qquad (30)$$

$$\Delta_j(f)(t) = \frac{1}{c_\psi}\int D_f(b,2^{-j})\Psi_{(b,2^{-j})}(t)db \qquad (31)$$

we obtain the following decomposition: for all $f \in L^2(\mathbb{R})$,

$$f = \Sigma_{j_0}(f) + \sum_{j=j_0+1}^{\infty} \Delta_j(f) = \sum_{j=-\infty}^{\infty} \Delta_j(f) \tag{32}$$

Such decompositions are called *Littlewood-Paley* decompositions, and play a great role in several contexts, for example in the study of partial differential equations. They are also extremely useful in the context of signal analysis, for they provide a version of wavelet analysis in which the translation invariance properties of the continuous transform are preserved. Littlewood-Paley decompositions may also be thought of as a first step towards multiresolution decompositions, which will be described in Section 4

3 Wavelets and the characterization of regularity

The wavelet transform is often compared to a *mathematical microscope*, because of its capability of zooming in. In fact, another crucial property of wavelets is their blindness to certain "regular" behaviors. Indeed, we have seen already that a wavelet has necessarily zero integral. Hence, two functions which differ only by a constant have the same wavelet coefficients: wavelets are blind to constants. If necessary, one may make them blind to higher order polynomial behavior, by imposing vanishing moments conditions, of the type

$$\int t^m \psi(t) dt = 0 \ , \quad m = 0, 1, \dots M \ .$$

This is best illustrated by the study of Hölder regularity properties of functions, which we sketch below.

3.1 Global regularity properties

We start with an example of characterization of some global regularity properties. Let us first recall some definitions. For $0 < \alpha < 1$, let C^α denote the Banach space of functions f satisfying

$$|f(t) - f(s)| \leq C|t-s|^\alpha, \qquad t, s \in \mathbb{R}, \tag{33}$$

for some constant $0 < C < \infty$. Such functions are said to be Hölder continuous of order α. Then we have

Theorem 3.1 *Let $\psi \in L^1(\mathbb{R})$ be such that*

$$c_\alpha = \int |t|^\alpha \, |\psi(t)| \, dt < \infty \ , \textit{and} \int \sup_{|\delta|<1} |\psi'(u+\delta)| \, du < \infty \ . \tag{34}$$

Then $f \in C^\alpha$ if and only if

$$|T_f(b,a)| \leq C' a^\alpha, \qquad a \in \mathbb{R}, \tag{35}$$

for some finite positive constant C' independent of b.

We will not give a complete proof of this result here (see, e.g., [26]), but in order to illustrate the role of the vanishing moments, let us sketch the proof of the first half. Let ψ be a wavelet satisfying the above properties, and let $f \in C^\alpha$. Then since $\int \psi(t)dt = 0$, we have $T_f(b,a) = \frac{1}{a}\int [f(t) - f(b)]\, \overline{\psi}\left(\frac{t-b}{a}\right) dt$, and

$$|T_f(b,a)| \leq \frac{C}{a}\int |t-b|^\alpha \left|\psi\left(\frac{t-b}{a}\right)\right| dt \leq C c_\alpha\ a^\alpha\ .$$

which proves the first part of the theorem. The converse is proved in a similar way and makes use of the second assumption on the wavelet.

Remark 3.1 If one associates to any $f \in C^\alpha$ the infimum $\|f\|_\alpha$ of the constants C such that (33) holds, it may be proved that $\|\cdot\|_\alpha$ provides C^α with a norm, which makes it a Banach space (notice that functions in this Banach space, are only defined modulo additive constants). Then Theorem 3.1 expresses the fact that for a suitably chosen wavelet, the number $\inf a^{-\alpha}|T_f(b,a)|$ defines an equivalent norm on C^α. This is one of the simplest examples of the characterization of function spaces by mean of the wavelet transform. Elaborating on such arguments leads to the characterization of wider classes of functional spaces (see, e.g., [46] or [26] for a review).

The theorem above only addresses the case of Hölder spaces of exponent $0 < \alpha < 1$. More generally, let us assume that $n < \alpha < n+1$. Then the function f belongs to the Hölder space C^α if for any t there exists a polynomial P_n of order n and a constant C such that for $|h|$ small enough,

$$|f(t) - P_n(h)| \leq C|h|^\alpha\ . \tag{36}$$

Equivalently, f is C^α if its derivative of order n is $C^{\alpha-n}$. To adapt to this new situation, extra assumptions on the wavelet are needed. To see this, let us consider the following toy example. Assume that f has $M-1$ continuous derivatives in a neighborhood U of a point $t = t_0$. Then for every $t \in U$ we may write

$$f(t) = f(t_0) + (t-t_0)f'(t_0) + \cdots + \frac{1}{(M-1)!}(t-t_0)^{M-1}f^{(M-1)}(t_0) + r(t)\ .$$

If the wavelet ψ (assumed to have compact support for the sake of the present argument) has M vanishing moments, then for sufficiently small values of the scale parameter a (i.e., such that $\mathrm{supp}(\psi_{(b,a)}) \subset U$ for some b), we have

$$T_f(b,a) = \langle f, \psi_{(b,a)}\rangle = \langle r, \psi_{(b,a)}\rangle\ .$$

Then we can say that, at least in some sense, the wavelet does not "see" the regular behavior of $f(t)$ near $t = t_0$ and "focuses" on the potentially singular part $r(t)$. If in addition we assume that $f(t)$ is C^M near $t = t_0$, then we have $r(t) = (t-t_0)^M\rho(t)$, with $|\rho(t)| \leq \sup|f^{(M)}|/M!$, and we immediately conclude that

$$T_f(b,a) \sim O(a^M) \quad \text{as} \quad a \to 0\ .$$

More generally, Theorem 3.1 is still valid if one assumes in addition that the wavelet ψ has n vanishing moments (and a corresponding generalization of condition (34)). The proofs go along the same lines, except that instead of using the zero integral property of ψ to subtract $f(b)$ from $f(t)$, one uses all the vanishing moment assumption to subtract $P_n(t-b)$ from $f(t)$ inside the integral.

Elaborating on such arguments leads to the characterization of Hölder spaces of functions via wavelet transform (see, e.g., [46, 26]) and the whole body of work on the characterization of singularities of a given signal [43]. More elaborate arguments also yield characterizations of more general functional classes, such as Sobolev, Besov and Triebel classes.

3.2 Local regularity

The previous results are global in the sense that they require uniform regularity throughout the real line. Another classical example involves local regularity. For a given point t we say that the function f is locally Hölder continuous of order α at t if it satisfies

$$|f(t) - f(s)| \leq C|t-s|^\alpha \tag{37}$$

for some constant C and for $|t-s|$ small enough. Then we have the "local" counterpart of the previous result:

Theorem 3.2 *Let $\psi \in L^1(\mathbb{R})$ be a wavelet with $M+1$ vanishing moments such that for any m, there is a constant C_m such that $\left|\psi^{(k)}(t)\right| \leq \frac{C_m}{1+|t|^m}$.*

1. *Let the function f be locally Hölder continuous of order α at t_0. Then*
$$|T_f(b,a)| \leq C'(a^\alpha + |b-t_0|^\alpha) \tag{38}$$
for some constant C'.

2. *Conversely, Let $\alpha < M$ be a non-integer number, and assume that there exist numbers $\alpha' < \alpha$ and A such that for all a, b*
$$|T(b,a)| \leq A a^\alpha \left(1 + \left|\frac{t_0-b}{a}\right|^{\alpha'}\right) . \tag{39}$$
Then f is locally Hölder continuous of order α at t_0.

The proof of the theorem uses the ingredients we have developed above. Let us start with the first part, and consider $T(b,a) = \frac{1}{a}\int (f(t) - P(t))\,\overline{\psi}\left(\frac{t-b}{a}\right) dt$.

$$\begin{aligned}
|T(b,a)| &\leq \frac{A}{a}\int |t-t_0|^\alpha \left|\psi\left(\frac{t-b}{a}\right)\right| dt \\
&\leq 2^\alpha \frac{A}{a}\int (|t-b|^\alpha + |b-t_0|^\alpha) \left|\psi\left(\frac{t-b}{a}\right)\right| dt \\
&\leq 2^\alpha A \left(a^\alpha \int |u|^\alpha |\psi(u)| du + |t_0 - b|^\alpha ||\psi||_1\right) ,
\end{aligned}$$

which proves the first assertion.

For the converse assertion, one first needs to exhibit the polynomial $P(t)$. Let us consider the Littlewood-Paley blocks

$$\Delta_j(t) = \int_{-\infty}^{\infty}\int_{2^j}^{2^{j+1}} T(b,a)\psi_{(b,a)}(t)\frac{da}{a}\,db$$

According to the assumptions made on ψ, each block may be bounded as

$$|\Delta_j(t)| \leq K 2^{j\alpha} \int_{-\infty}^{\infty} \frac{1 + 2^{\alpha'}|u|^{\alpha'} + 2^{\alpha'}\left|\frac{t-t_0}{2^j}\right|^{\alpha'}}{1+|u|^m}\, du \ .$$

Taking $m > \alpha' + 1$ yields the following estimate:

Lemma 3.1

$$|\Delta_j(t)| \leq K 2^{j\alpha} \left(1 + \left|\frac{t-t_0}{2^j}\right|^{\alpha'}\right) \ . \tag{40}$$

In the same way, we obtain the estimates

Lemma 3.2 *For all k smaller than the integer part of α,*

$$\left|\Delta_j^{(k)}(t)\right| \leq K 2^{j(\alpha-k)} \left(1 + \left|\frac{t-t_0}{2^j}\right|^{\alpha'}\right) \ . \tag{41}$$

Let us now consider the truncated Taylor expansion

$$P_{t_0}(t) = \sum_{k=0}^{\lfloor\alpha\rfloor} \left(\sum_{j=-\infty}^{\infty} \Delta_j^{(k)}(t_0) \right) \frac{(t-t_0)^k}{k!}$$

The behavior of this sum as $j \to -\infty$ is easily controlled thanks to Lemma 3.2. In addition, the fact that $|T(b,a)| \leq ||f||\,||\psi||/\sqrt{a}$ implies $|\Delta_j^{(k)}(t_0)| \leq K 2^{-j(k+\frac{1}{2})}$, which takes care of the limit $j \to \infty$. Then, $P_{t_0}(t)$ is well defined. Consider now $f(t) - P_{t_0}(t)$. Let J be such that $2^{J-1} \leq |t-t_0| < 2^J$. It follows from Taylor's formula that

$$\sum_{j=J}^{\infty} \left[\Delta_j(t) - \sum_{k=0}^{\lfloor\alpha\rfloor} \Delta_j^{(k)}(t_0) \frac{(t-t_0)^k}{k!} \right] \leq \frac{|t-t_0|^{\lfloor\alpha\rfloor+1}}{(\lfloor\alpha\rfloor+1)!} \sup_{s\in[t,t_0]} \left|\Delta_j^{(\lfloor\alpha\rfloor+1)}(s)\right| \leq K|t-t_0|^\alpha$$

In addition, the sum from $-\infty$ to $J-1$ is bounded by

$$\begin{aligned}
& K \sum_{j=-\infty}^{J-1} \left[2^{j\alpha} \left(1 + \left|\frac{t-t_0}{2^j}\right|^{\alpha'}\right) + \sum_{k=0}^{\lfloor\alpha\rfloor} 2^{j(\alpha-k)} \frac{|t-t_0|^k}{k!} \right] \\
\leq \quad & K \left[2^{J\alpha} + 2^{J(\alpha-\alpha')}|t-t_0|^{\alpha'} + \sum_{k=0}^{\lfloor\alpha\rfloor} 2^{J(\alpha-k)} \frac{|t-t_0|^k}{k!} \right] \\
\leq \quad & K \left[2^\alpha |t-t_0|^\alpha + 2^{\alpha-\alpha'}|t-t_0|^\alpha + \sum_{k=0}^{\lfloor\alpha\rfloor} \frac{2^{\alpha-k}}{k!} |t-t_0|^\alpha \right] \leq K|t-t_0|^\alpha \ ,
\end{aligned}$$

where K is a constant, which may change from line to line. This sketches the proof. More details may be found for example in [37] and [43] which contain reviews especially tailored to mathematicians and signal analysts, respectively.

Remark 3.2 It is instructive to examine the meaning of conditions of the type (38). Let us consider the cone in the (b, a) half-plane defined by the condition $|b - t_0| < a$. Within this cone we have $|T_f(b, a)| = O(a^\alpha)$ as $a \to 0$. Outside the cone, the behavior is governed by the distance of b to the point t_0. These two behaviors are generally different, and have to be studied independently. However, it is shown in [43] that non-oscillating singularities may be characterized by the behavior of their wavelet transform within the cone.

Remark 3.3 It is also shown in [43] that rapidly oscillating singularities cannot be characterized by the behavior of their wavelet transform in the cone. A typical example is given by the function $f(t) = \sin 1/t$ whose instantaneous frequency tends to infinity as $x \to 0$. The wavelet transform modulus is maximum on a curve of equation $b = Ka^2$ for some constant K depending only on the wavelet, and this curve is not in the cone. In such a case, the oscillations have to be analyzed carefully (see, e.g., [37].)

Remark 3.4 We close this discussion with a more "practical" remark. Most often, data are only available in the form of discrete signals, i.e. only for discrete values of the variable (with a fixed resolution given by the sampling frequency). Therefore, the notions of singularities and Hölder exponents are strictly speaking meaningless (as is the limit $a \to 0$.) Nevertheless, one can say that the behavior of the wavelet coefficients across scales provides a good way of describing the regularity of functions whose samples coincide with the observations up to a given resolution.

4 Multiresolution wavelet analysis

Let us now turn to an approach which is completely different, namely the multiresolution approach. This approach is based upon a general methodology (the *multiresolution analysis*, which automatically generates orthonormal bases of wavelets. Remarkably enough, such an approach turned out to be extremely close to image processing methods (subband filtering methods), which led to fast pyramidal algorithms for wavelet decompositions.

We shall not go into the details of the theory, which received a considerable attention in the recent years, and refer to [15, 53, 55] for details. We rather sketch the construction and give a few consequences and examples.

4.1 Multiresolution analysis and associated wavelet basis

Let us start by introducing the notion of multiresolution analysis.

Definition 4.1 *A multiresolution analysis (MRA) of $L^2(\mathbb{R})$ is a collection of nested closed subspaces $V_j \subset L^2(\mathbb{R})$*

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \ldots \tag{42}$$

such that the following properties hold:

1. $\overline{\cup V_j} = L^2(\mathbb{R})$ *and* $\cap V_j = \{0\}$.
2. *If $f \in V_0$, then $f(\cdot - k) \in V_0$ for all $k \in \mathbb{Z}$; $f \in V_j$ if and only if $f(\cdot/2) \in V_{j-1}$.*

3. *There exists a function $\chi \in V_0$ such that the collection of the integer translates $\chi(\cdot - k)$ for $k \in \mathbb{Z}$ is a Riesz basis of V_0.*

The function χ is called a *scaling function*, and the Riesz basis may be orthonormalized by Gram's procedure by setting

$$\hat{\phi}(\omega) = \frac{\hat{\chi}(\omega)}{\sqrt{\sum_k |\hat{\chi}(\omega + 2\pi k)|^2}} \ . \tag{43}$$

Corollary 4.1 *With the same notation as before, the collection $\{\phi(\cdot - k);\, k \in \mathbb{Z}\}$ is an orthonormal basis of V_0.*

The function $\phi(t)$ is also called a scaling function. An immediate consequence of the definition of ϕ in (43) is that the Fourier transform satisfies

$$\sum_k |\hat{\phi}(\omega + 2\pi k)|^2 = 1 \ . \tag{44}$$

From now on, we shall essentially work with the scaling function ϕ associated with orthonormal basis. The inclusion of the V_j spaces shows that $\phi(t)$ (as well as $\chi(t)$) may be expressed as a linear combination of the functions $2^{1/2}\phi(2t-k)$, which form an orthonormal basis of V_1. The consequence is the so-called *two-scale difference equation* (or *refinement equation*)

$$\phi(t) = \sqrt{2}\sum_k h_k \phi(2t + k) \ . \tag{45}$$

Let now the 2π-periodic function $H(\omega)$ be defined by

$$H(\omega) = \frac{1}{\sqrt{2}}\sum_k h_k e^{ik\omega} \ , \tag{46}$$

Clearly, we have $\hat{\phi}(2\omega) = H(\omega)\hat{\phi}(\omega)$. Let W_j denote the orthogonal complement of V_j in V_{j+1}. The cornerstone of the multiresolution theory is the existence of a function $\psi \in W_0$ such that the collection $\{\psi(\cdot - k), k \in \mathbb{Z}\}$ is an orthonormal basis of W_0. More precisely, if we define the 2π-periodic function $G(\omega)$ by

$$G(\omega) = e^{i\omega}\overline{H(\omega + \pi)}, \tag{47}$$

and if we denote by $2^{-1/2}g_k$ its Fourier coefficients, i.e., if we set

$$\begin{aligned} G(\omega) &= \ \tfrac{1}{\sqrt{2}}\textstyle\sum_k g_k e^{ik\omega}, \\ g_k &= \ -(-1)^k \overline{h}_{1-k} \ , \quad k \in \mathbb{Z} \ , \end{aligned} \tag{48}$$

then we have

Theorem 4.1 *If $\psi(t)$ is defined by*

$$\psi(t) = \sqrt{2}\sum_k g_k \phi(2t + k), \qquad\qquad t \in \mathbb{R}, \tag{49}$$

then the collection $\{\psi(\cdot - k);\, k \in \mathbb{Z}\}$ is an orthonormal basis of W_0.

See, e.g., [46] for a proof. Notice that ψ constructed in this way is in fact an analyzing wavelet in the sense of the continuous wavelet transform introduced earlier. Formula (49) is a second example of refinement equation. Set

$$\left\{ \begin{array}{rcl} \psi_{jk}(t) & = & 2^{j/2}\psi\left(2^j t - k\right) \ , \\ \phi_{jk}(t) & = & 2^{j/2}\phi\left(2^j t - k\right) \ . \end{array} \right. \tag{50}$$

By definition, it follows that for a given $j \in \mathbb{Z}$, the collection $\{\phi_{jk};\ k \in \mathbb{Z}\}$ (resp. $\{\psi_{jk};\ k \in \mathbb{Z}\}$) is an orthonormal basis of V_j (resp. W_j), and we have

Corollary 4.2 *With the same notation as before, the collection of wavelets $\{\psi_{jk};\ j,k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(\mathbb{R})$.*

4.2 Quadrature mirror filters

The remarkable property of $H(\omega)$ and $G(\omega)$ is that they satisfy the so-called *Quadrature Mirror Filter* (QMF) condition, which is a direct consequence of equations (44) and (45):

Lemma 4.1 *The periodic functions $H(\omega)$ and $G(\omega)$ defined in* (46) *and* (47) *satisfy*

$$|H(\omega)|^2 + |G(\omega)|^2 = 1, \qquad \omega \in \mathbb{R}. \tag{51}$$

Let now $f \in L^2(\mathbb{R})$ be any square-integrable function, and set

$$s_{jk} = \langle f, \phi_{jk} \rangle \ , \quad \text{and} \quad d_{jk} = \langle f, \psi_{jk} \rangle \tag{52}$$

Then, an immediate consequence of the refinement equations and Lemma 4.1 is

Lemma 4.2

1. *The coefficients $\{d_{jk,j,j\in\mathbb{Z}}\}$ and $\{s_{jk,j,j\in\mathbb{Z}}\}$ may be computed recursively via the following pyramidal algorithm*

$$s_{j-1\,k} = \sum_\ell \overline{h}_{2k-\ell} s_{j\ell} \ , \quad \textit{and} \quad d_{j-1\,k} = \sum_\ell \overline{g}_{2k-\ell} s_{j\ell} \ . \tag{53}$$

2. *The finer resolution coefficients may be obtained from the coarser ones by*

$$s_{j+1\,\ell} = \sum_k \left(h_{2k-\ell} s_{j\ell} + g_{2k-\ell} d_{j\ell}\right) \ . \tag{54}$$

These two equations are of fundamental practical importance, for they automatically yield fast algorithms for discrete wavelet transform and inverse transform.

Let us also mention that such pyramidal algorithms have been generalized to yield new orthonormal bases of $L^2(\mathbb{R})$, known as wavelet packet bases. The interested reader may refer to [55] for a detailed account of such generalizations.

4.3 Discrete Littlewood-Paley transform

Let us finally quote for the record a variant of discrete wavelet transform, which is essentially a discrete version of the Littlewood-Paley transform we described above. The goal of such a transform is to restore a kind of (discrete) translation invariance. Let us consider a multiresolution analysis, with scaling function $\phi(t)$ and wavelet $\psi(t)$, and introduce the following family of functions, defined by

$$\left\{ \begin{array}{lcl} \psi_k^j(t) & = & 2^{j/2}\psi\left(2^j(t-k)\right) \ , \\ \phi_k^j(t) & = & 2^{j/2}\phi\left(2^j(t-k)\right) \ . \end{array} \right. \tag{55}$$

This family forms an overcomplete set. For any $f \in L^2(\mathbb{R})$, define

$$D_k^j = \langle f, \psi_k^j \rangle \ , \quad \text{and} \quad S_k^j = \langle f, \phi_k^j \rangle \ . \tag{56}$$

An immediate calculation yields the following pyramidal algorithm for computing such coefficients:

$$S_k^{j-1} = \sum_\ell \overline{h}_{k-2^{j-1}\ell} S_\ell^j \ , \quad \text{and} \quad D_k^{j-1} = \sum_\ell \overline{g}_{k-2^{j-1}\ell} S_\ell^j \ . \tag{57}$$

These coefficients yield an alternate wavelet representation for the function $f(t)$, suitable for problems in which translation invariance is an important issue. See for example [43] for some applications in the image processing context.

5 Implementing Symmetries

As we have seen already, the continuous (linear or quadratic) time-frequency transforms described at the beginning of this paper possess important built-in symmetry properties. Such covariance properties, which are crucial for practical purposes (in particular for signal processing applications) are in fact intrinsic to the transforms, and even characterize them. This fact is conveniently described by means of group-theoretical methods. Indeed, the set of simple transformations used to generate the wavelets from a single one in general inherits the structure of a group G (as is the case for instance for translations, modulations or dilations).

The deep connection between the usual wavelet decompositions, coherent states theory and the theory of square-integrable group representations was emphasized by H. Moscovici and A. Verona [47, 48] and A. Grossmann, J. Morlet and T. Paul independently [30]. As a result, all such wavelets are generated from a representation of a group of simple transformations G, in such a way that the representation is unitarily equivalent to a subrepresentation of the regular representation, i.e. a representation of the group G onto $L^2(G)$.

5.1 Square-integrable group representations

The connexion between time-frequency representation theorems and square-integrable group representations was realized by A. Grossmann, J. Morlet and T. Paul in [30]. We briefly sketch the construction.

Definition 5.1 *Let G be a separable locally compact Lie group, and let π be a unitary strongly continuous representation of G on the Hilbert space $\mathcal{H}$. π is said to be square-integrable (or to belong to the discrete series of G) if π is irreducible, and if there exists at least a vector $v \in \mathcal{H}$ such that*

$$0 \ < \ \int_G |\langle \pi(g)v, v\rangle|^2 \, d\mu(g) \ < \ \infty \ . \tag{58}$$

Such a vector is said to be admissible.

Square-integrable group representations have been extensively studied in the literature, in particular for compact groups, locally compact unimodular groups and non-unimodular locally compact groups. The results may be summarized in the following theorem, due to Duflo and Moore:

Theorem 5.1 *Let π be a square-integrable strongly continuous unitary representation of the locally compact group G on $\mathcal{H}$. Then there exists a positive self-adjoint operator C such that for any admissible vectors $v_1, v_2 \in \mathcal{H}$ and for any $u_1, u_2 \in \mathcal{H}$*

$$\int_G \langle u_1, \pi(g).v_1\rangle\langle\pi(g).v_2, u_2\rangle \, d\mu(g) = \langle C^{1/2}v_2, C^{1/2}v_1\rangle \ \langle u_1, u_2\rangle \tag{59}$$

Moreover, the set of admissible vectors coincides with the domain of C.

Let $\lambda : G \to \mathcal{U}(L^2(G))$ be the left-regular representation of G: if $f \in L^2(G)$,

$$(\lambda(h)f)(g) = f(h^{-1}g) \ . \tag{60}$$

Theorem 5.1 shows that a representation π of G is square integrable if and only if is is unitarily equivalent to a subrepresentation of the left-regular representation λ. The corresponding intertwiners can be realized as follows. If v is an admissible vector in $\mathcal{H}$, and $v' \in \mathcal{H}$, introduce the Schur coefficients, i.e the matrix coefficients of elements of G:

$$c_{v,v'}(g) = \langle v', \pi(g)v\rangle, \ g \in G \tag{61}$$

Let T be the left transform [30], i.e. the map defined by

$$T : u \in \mathcal{H} \to T_u = c_{v,u}(\cdot) \in L^2(G) \tag{62}$$

T realizes the intertwinning between π and λ:

$$T \circ \pi = \lambda \circ T \tag{63}$$

The idea of Grossmann, Morlet and Paul was to use (59) and (62) for the analysis of functions, in the case where $\mathcal{H}$ is a function space. This was the starting point of many applications, especially in a signal analysis context. The left transform T is used to obtain another representation of functions, and (63) expresses the covariance of the transform. Notice that the continuity assumption is fulfilled by construction.

5.2 The classical examples

Gabor functions We start with the case of the so-called *canonical coherent states*, generated from the n-dimensional *Weyl-Heisenberg group* $G_{WH} = \mathbb{R}^{2n} \times S^1$, with group operation

$$(\underline{q}, \underline{p}, \varphi) \cdot (\underline{q}', \underline{p}', \varphi') = (\underline{q} + \underline{q}', \underline{p} + \underline{p}', \varphi + \varphi' + \underline{p}.\underline{q}' \ [mod\, 2\pi]) \tag{64}$$

It follows from the Stone-Von-Neumann theorem that any irreducible unitary representations of G_{WH} is unitarily equivalent to one of the following form (see [40, 52]). If $f \in L^2(\mathbb{R}^n)$

$$\left[\pi(\underline{q}, \underline{p}, \varphi) \cdot f\right](\underline{x}) = e^{i\mu(\varphi - \underline{p}.(\underline{x} - \underline{q}))} f(\underline{x} - \underline{q})$$

for some $\mu \in \mathbb{Z}^*$. Let us focus on the one-dimensional situation for the sake of simplicity, and assume $\mu = 1$. The representation is square-integrable, and the corresponding left transform is the Gabor transform described in Definition 2.3, up to a phase factor $e^{i\varphi}$. More precisely, the Gaborlets read

$$g_{(b,\omega)} = \pi(b, \omega, 0)g \ .$$

This can be interpreted as a square-integrable projective representation of the abelian group $\mathbb{R}^{2n}$. Notice that in such a case, one does not have full covariance by time-frequency shifts, precisely because of this phase factor (twisted covariance.)

One-dimensional wavelets Let $G_{aff} = \mathbb{R} \times \mathbb{R}_+^*$ denote the one-dimensional *affine group*, or "$ax + b$" *group* with group operation

$$(b, a) \cdot (b', a') = (b + ab', aa')$$

It may be proved G_{aff} has only two inequivalent irreducible unitary representations on $H^2_\pm(\mathbb{R}) = \{f \in L^2(\mathbb{R}), \hat{f}(\pm\omega) = 0 \, \forall \omega \geq 0\}$, of the form

$$[\pi(b, a) \cdot f] (t) = \frac{1}{\sqrt{a}} f\left(\frac{t - b}{a}\right) \tag{65}$$

(notice the slight change in the normalization.) This representation is square-integrable, and the corresponding left transform is nothing but the affine wavelet transform described in Definition 2.5. The covariance equation (63) simply expresses that the affine wavelet transform of a dilated and shifted copy of a function $f(t)$ coincides with a dilated and shifted version of the wavelet transform of $f(t)$. Finally, the admissibility of a vector $\psi \in L^2(\mathbb{R})$ reduces to (21). In a signal analysis context, the translation parameter is interpreted as a position (or time) parameter, and the scale parameter a as a frequency parameter (more precisely the inverse of a frequency parameter). One is then led to a *time-frequency representation*.

***N*-dimensional wavelets** To generalize the previous construction to the n-dimensional case, one faces an irreducibility problem. The n-dimensional affine group $G_{aff} = \mathbb{R}^n \times \mathbb{R}_+^*$ does not act irreducibly on $L^2(\mathbb{R}^n)$. To restore irreducibility, R. Murenzi [49] proposed to consider the semi-direct product of the affine group by $SO(n)$. The resulting group, denoted by $IG(n)$ then yields wavelets of the form

$$\psi_{(\underline{b},a,\underline{\underline{r}})}(\underline{x}) = a^{-n/2} \, \psi\left(r^{-1} \frac{\underline{x} - \underline{b}}{a}\right) \ , \tag{66}$$

and the following resolution of the identity: for all $f \in L^2(\mathbb{R}^n)$

$$f = C_\psi \int \langle f, \psi_{(\underline{b},a,\underline{\underline{r}})} \rangle \, \psi_{(\underline{b},a,\underline{\underline{r}})} \, \frac{d\underline{b}}{a^n} \, \frac{da}{a} \, d\theta \ . \tag{67}$$

5.3 Generalizations

Several extensions and generalizations have been developed. A first way of generalization concerns the study of more general classes of groups possessing square-integrable representations. Let us mention for example the works of Bernier and Taylor [5], Führ [27] and Aniello et al [2] on semi-direct products of the form $H \times \mathbb{R}^n$, where H is a closed subgroup of $GL(n, \mathbb{R})$, with group law $(h, \underline{b})(h', \underline{b}') = (hh', \underline{b} + h\underline{b}')$. The above mentioned authors have in particular given necessary and sufficient conditions on the group H which ensure that the representation under consideration is square integrable.

The application of the general theory to the case of discrete groups has also been studied. We refer to [24, 3] for more details.

In more general situations, i.e. for more general groups, the general theory does not apply, since the representation under consideration is not square-integrable. However, a representation may become square-integrable when restricted to an appropriate quotient group. The general theory has been developed by Antoine, Ali and Gazeau, and may be found in [1], where several particular cases are also studied.

6 Conclusion

We have described in this paper a (limited) number of aspects of wavelet analysis and time-frequency analysis, with emphasis on some signal processing related problems. Our goal was to give an idea of the broadness of the theory, as well as the huge range of applications and potential applications. The interested reader may refer to several excellent textbooks available in the literature (up to now, more than a hundred monographs have been published on the subject).

Before closing this review, let us stress that there are still many problems, connected to wavelet theory and applications, which are still to be solved. For the record, let us mention a few of them.

First, the connection between continuous and discrete wavelet systems is not completely understood. We have seen the algebraic and geometrical origin of continuous wavelet decompositions. The multiresolution approach seems to be also extremely constrained by algebraic arguments, which should be developed further (see [3]).

Also, a very important problem (from the practical point of view) is that of adaptive decomposition. We have seen that the "time-frequency toolbox" offers several different ways of decomposing a signal, which raises the problem of selecting the most appropriate one. This point has been studied by several authors (see for example [55]), and should be one of the main research areas in the near future.

References

[1] S.T. Ali, J.P. Antoine, J.P. Gazeau, and U.A. Müller (1995): Coherent states and their generalizations: an overview. *Rev. Math. Phys.* **7**, 1013–1104.

[2] P. Aniello, G. Cassinelli, E. DeVito and A. Levrero (1998): Wavelet transforms and discrete frames associated to semidirect products. *J. Math. Phys.* **39**, 3965-3973.

[3] J.P. Antoine, Y. Kougagou, D. Lambert and B. Torrésani (1998): An algebraic approach to discrete dilations. Submitted to *J. of Fourier Analysis and Applications*.

[4] E. Bacry, J.F. Muzy, and A. Arneodo (1993): Singularity spectrum of fractal signals from wavelet transform: exact results. *J. Stat. Phys.* **70**, 635–674.

[5] D. Bernier and M. Taylor (1996): Wavelets and square integrable representations. *SIAM J. Math. Anal.* **27**, 594–608

[6] P. Bertrand and J. Bertrand (1985): Représentation temps-fréquence des signaux à large bande. *La Recherche Aérospatiale (in French)* **5**, 277–283

[7] G. Beylkin, R. Coifman, and V. Rokhlin (1991): Fast wavelet transforms and numerical algorithms I. *Comm. Pure Appl. Math.* **44**, 141–183.

[8] R.E. Blahut, W. Miller Jr, and C.H. Wilcox, Eds. (1991): Radar and Sonar, Part I. *Springer Verlag*, New York, N.Y.

[9] A. Calderón (1964): Intermediate spaces and interpolation, the complex method. *Studia Math.* **24**, 113

[10] S. Cambanis and C. Houdré (1995): On continuous wavelet transforms of second order random processes. *IEEE Trans. Inf. Theory* **41**, 628–642.

[11] R. Carmona, W.L. Hwang and B. Torrésani, *Practical Time-Frequency Analysis: Gabor and wavelet transforms, with an implementation in S*, Academic Press (1998).

[12] C.K. Chui (1992): *An Introduction to Wavelets*, Academic Press.

[13] A. Cohen, I. Daubechies, and J.C. Feauveau (1992): Biorthogonal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.* **45**, 485–560.

[14] R.R. Coifman and Y.R. Meyer (1991): Remarques sur l'analyse de Fourier à fenêtre. *C.R. Acad. Sci. Paris Série I Math.*, **312**, 259–261.

[15] I. Daubechies (1992): *Ten Lectures on Wavelets*. Vol. 61, CBMS-NFS Regional Series in Applied Mathematics.

[16] I. Daubechies (1988): Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* **41**, 909–996.

[17] I. Daubechies (1993): Orthonormal bases of compactly supported wavelets II. Variations on a theme, *SIAM J. Math. Anal.* **24**, 499–519.

[18] I. Daubechies, S. Jaffard, and J.L. Journé (1991): A simple Wilson basis with exponential decay. *SIAM J. Math. An.***22**, 554–572.

[19] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian, and B. Torrésani (1992): Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies. *IEEE Trans. Inf. Th.* **38**, 644–664.

[20] D. Donoho, I. Johnstone, G. Kerkyacharian, and D. Picard (1995): Wavelet shrinkage: asymptotia ? *J. Royal Stat. Soc.* **57**, 301–337.

[21] M. Duflo, C.C. Moore (1976): On the regular representation of a nonunimodular locally compact group, *J. Funct. An.* **21**, 209-243.

[22] M. Duval-Destin, M.A. Muschietti, and B. Torrésani (1993): Continuous wavelet decompositions: multiresolution and contrast analysis. *SIAM J. Math. An.* **24**, 739–755.

[23] P. Flandrin (1993): Temps-Fréquence. *Traité des Nouvelles Technologies, série Traitement du Signal, Hermes*; in French; english translation to appear at Academic Press (1999).

[24] K. Flornes, A. Grossmann, M. Holschneider, and B. Torrésani (1994): Wavelets on Discrete Fields. *Appl. and Comp. Harmonic Anal.* **1** 137–146.

[25] M. Folland: *Harmonic Analysis of Phase Space* Princeton University Press.

[26] M. Frazier, B. Jawerth, and G. Weiss (1991): *Littlewood-Paley Theory and the Classification of Function Spaces*, Regional Conference Series in Mathematics **79**, Providence, RI, American Mathematical Society.

[27] H. Führ (1996): Wavelet frames and admissibilityin higher dimensions. *J. Math. Phys.* **37**, 6353-6366.

[28] A. Grossmann, R. Kronland-Martinet, and J. Morlet (1987): Reading and understanding the continuous wavelet transform. In *Wavelets, Time-Frequency Methods and Phase Space*, J.M. Combes, A. Grossmann and Ph. Tchamitchian, Eds., IPTI, Springer Verlag, 2–20.

[29] A. Grossmann, J. Morlet (1984): Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. of Math. An.* **15**, 723–736.

[30] A. Grossmann, J. Morlet, and T. Paul (1985): Transforms associated with square integrable group representations I. *J. Math. Phys.* **27**, 2473–2479.

[31] A. Grossmann, J. Morlet, and T. Paul (1986): Transforms associated with square integrable group representations II. *Ann. Inst. H. Poincaré* **45**, 293–309.

[32] A. Guichardet (1985): Théorie de Mackey et méthode des orbites selon M. Duflo, *Expo. Math.* **3** 303-346.

[33] F. Hlawatsch and G. Matz (1998):

[34] J.A. Hogan and J.D. Lakey (1995): Extensions of the Heisenberg group by dilations and frames. *Appl. Comp. Harm. Anal* **2**, 174–199.

[35] M. Holschneider (1988): On the wavelet transformation of fractal objects. *J. Stat. Phys.* **50**, 953–993.

[36] M. Holschneider and Ph. Tchamitchian (1991): Pointwise analysis of Riemann-Weierstrass "nowhere differentiable" function. *Invent. Math.* **105**, 157–176.

[37] S. Jaffard and Y. Meyer (1996): Wavelet Methods for Pointwise Regularity and Local Oscillations of Functions. *Memoirs of the AMS* **587**.

[38] W. Kozek (1996): *Spectral Estimation in Non-Stationary Environments.* Ph.D. Thesis, Vienna.

[39] E.H. Lieb (1990): Integral bounds for radar ambiguity functions and Wigner distributions. *J. Math. Phys.* **31**, pp. 594-599.

[40] G. Mackey, Theory of unitary group representations, Univ. of Chicago Press (1976).

[41] S. Mallat (1998): *A Wavelet Tour of Signal Processing.* Academic Press, New York, N.Y.

[42] S. Mallat (1989): A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. PAMI* **11**, 674–693.

[43] S. Mallat and W.L. Hwang (1992): Singularities detection and processing with wavelets. *IEEE Trans. Info. Theory* **38#2**, 617–643.

[44] S. Mallat, G. Papanicolaou, and Z. Zhang (1995): Adaptive covariance estimation of locally stationary processes. *Ann. Stat.* **26**, 1–47.

[45] R.J. McAulay and T.F. Quatieri (1986): Speech analysis/synthesis based on a sinusoidal representation. *IEEE Trans. on Audio, Speech and Sign. Proc.* **34** 744–754.

[46] Y. Meyer (1989): *Ondelettes et opérateurs*,I: Ondelettes; II: Opérateurs de Calderón-Zygmund; III: (with R. Coifman) Opérateurs multilinéaires, Hermann. (English translation of first volume is published by Cambridge University Press.)

[47] H. Moscovici (1977): Coherent states representations of nilpotent Lie groups, *Comm. Math. Phys.* **54**, 63-68.

[48] H. Moscovici, A. Verona (1978): Coherent states and square-integrable representations, *Ann. I.H.P., Physique Théorique* **9**, 139-156.

[49] R. Murenzi (1990) Ondelettes multidimensionnelles et application à l'analyse d'images, PhD Thesis, Louvain-la-Neuve.

[50] M.B. Priestley (1965): Evolutionary Spectra and Non-Stationary Processes. *J. Roy. Stat. Soc.* **B27**, 204–237.

[51] R. Ryan and B. Torrésani, *Introduction to Continuous Wavelet Analysis*, SIAM, to appear (1999).

[52] W. Schempp, Harmonic analysis on the Heisenberg nilpotent Lie group, Pitman series 147 (1986), Wiley.

[53] M. Vetterli and J. Kovacevic (1996): *Wavelets and SubBand Coding*, Prentice Hall, Englewood Cliffs, NJ.

[54] J. Ville (1948): Théorie et applications, de la notion de signal analytique. *Cables et Transmissions* **2**, 61–74. Translated into English by I. Selin, RAND Corp. Report T-92, Santa Monica, CA (August 1958).

[55] M.V. Wickerhauser (1994): *Adapted Wavelet Analysis, from Theory to Software.* A.K. Peters Publ.

[56] E.P. Wigner (1932): On the quantum corrections for the thermodynamic equilibrium. *Phys. Rev.***40**, 749–759

[57] A. Zygmund (1959): *Trigonometric Series*, Cambridge University Press.

FIGURES

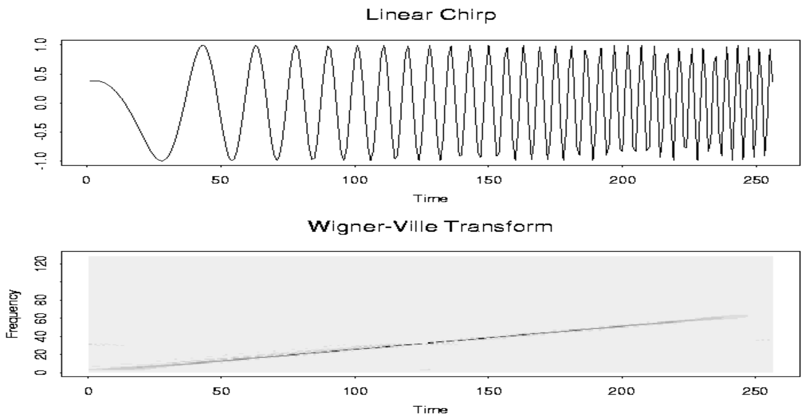

Figure 1: Example of a WV transform: the case of a linear chirp (coded with gray levels).

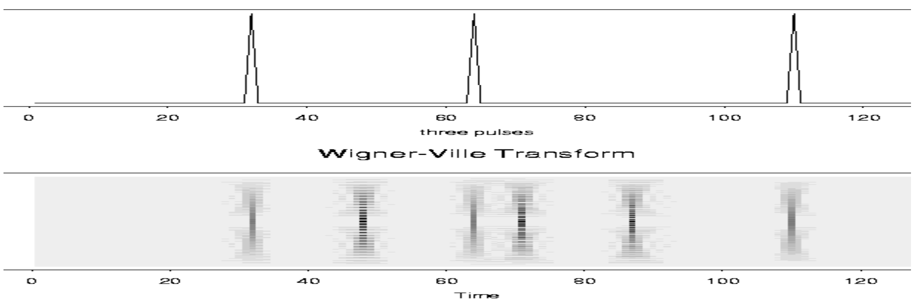

Figure 2: Interferences with the WV transform: the case of the sum of three pulses.

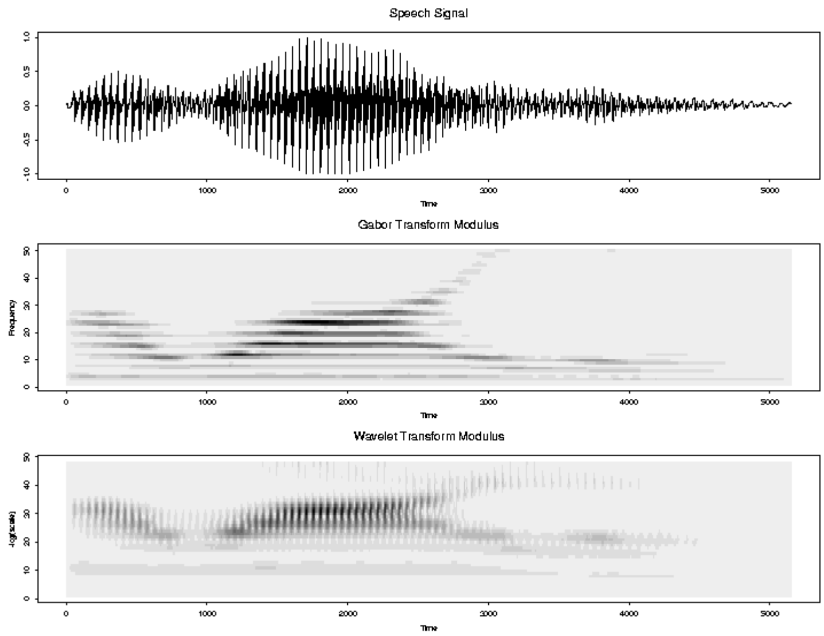

Figure 3: Gabor (middle) and wavelet (bottom) transforms of 625ms of speech signal (top).